1. For task one if we convert values to int and ignore decimal values, then 51 costs are calculated correctly. If we consider decimal values then only one cost is calculated correctly. On the other hand, the int part of the data is almost correctly calculated in the other rows and value of costs almost are near to each other once we compare the costs’ value of two tables api\_adwords\_costs and session\_sources. However, there are a few calculated costs that have a big difference between the two tables like campaign\_id\_79 and campaign\_id\_94. Moreover, there are no inconsistence rows (primary key duplication). Hence, I did not detect inconsistency via group by on user\_id, user\_id, event\_date and making condition having count (\*) > 1. Therefore, the count results for checking inconsistence rows are equal to null.
2. I could not find any pattern over time (I did various categorizations) and conversations are not stable overtime and they have high different revenue’s values as well. Hovewer, each unique value of conv\_id, user\_id, conv\_date, market has only one market.
3. The sum values of revenue’s column are identical. There is 16 conv\_id that two times came in the conv\_id column. Count all markets in both tables are identical. The highest conversation market value is GB and the lowest are EU, HR, MT.
4. There is no inconsistency because we do not have a duplicate on primary key values. Moreover, the **count** of rows that sum ihc less than one is countAttrLess1 19500. So, these rows of data don't make sense because the sum ihc of the is less than one. Also, the **count** of rows that sum ihc equal to one is 2993.
5. (Bonus 1) There are no special issues in the channel\_name column, considering I checked null values and special characters in this column. However, it should be checked why the sum CPC of some categories line Direct Traffic, Direct, Display, and Remarketing are zero. **(Second part)** No, the number of sessions is not stable per channels and they have different and big variations.
6. There is high number of duplicate rows in table session\_sources that I detected via group by key columns campaign\_id, user\_id, event\_date. So, there is inconsistency in this table. Moreover, the highest session number indicated for DE.